Exercise1

Load LinkedIn Connections

```
connections <- read.csv("/Users/moizshaikh/Downloads/Connections.csv", sep = ",", ski
p = 3)
# Use View(connections) in RStudio to manually inspect the data
```

Total Number of Connections

```
total_connections <- nrow(connections)
print(paste("Total number of connections:", total_connections))
```

```
## [1] "Total number of connections: 1078"
```

Data Analysis: Count Contacts by Company

```
company_counts <- connections %>%
  group_by(Company) %>%
  summarise(Count = n()) %>%
  arrange(desc(Count))
print(company_counts)
```

```
## # A tibble: 650 × 2
##    Company                                                Count
##    <chr>                                                  <int>
##  1 "Lahore University of Management Sciences"                40
##  2 ""                                                        30
##  3 "Desautels Capital Management"                            18
##  4 "McKinsey & Company"                                      18
##  5 "BAT"                                                     17
##  6 "Acasus - Brave Change"                                   14
##  7 "McGill University"                                       13
##  8 "McGill University - Desautels Faculty of Management"     13
##  9 "Unilever"                                                13
## 10 "TD"                                                      12
## # ℹ 640 more rows
```

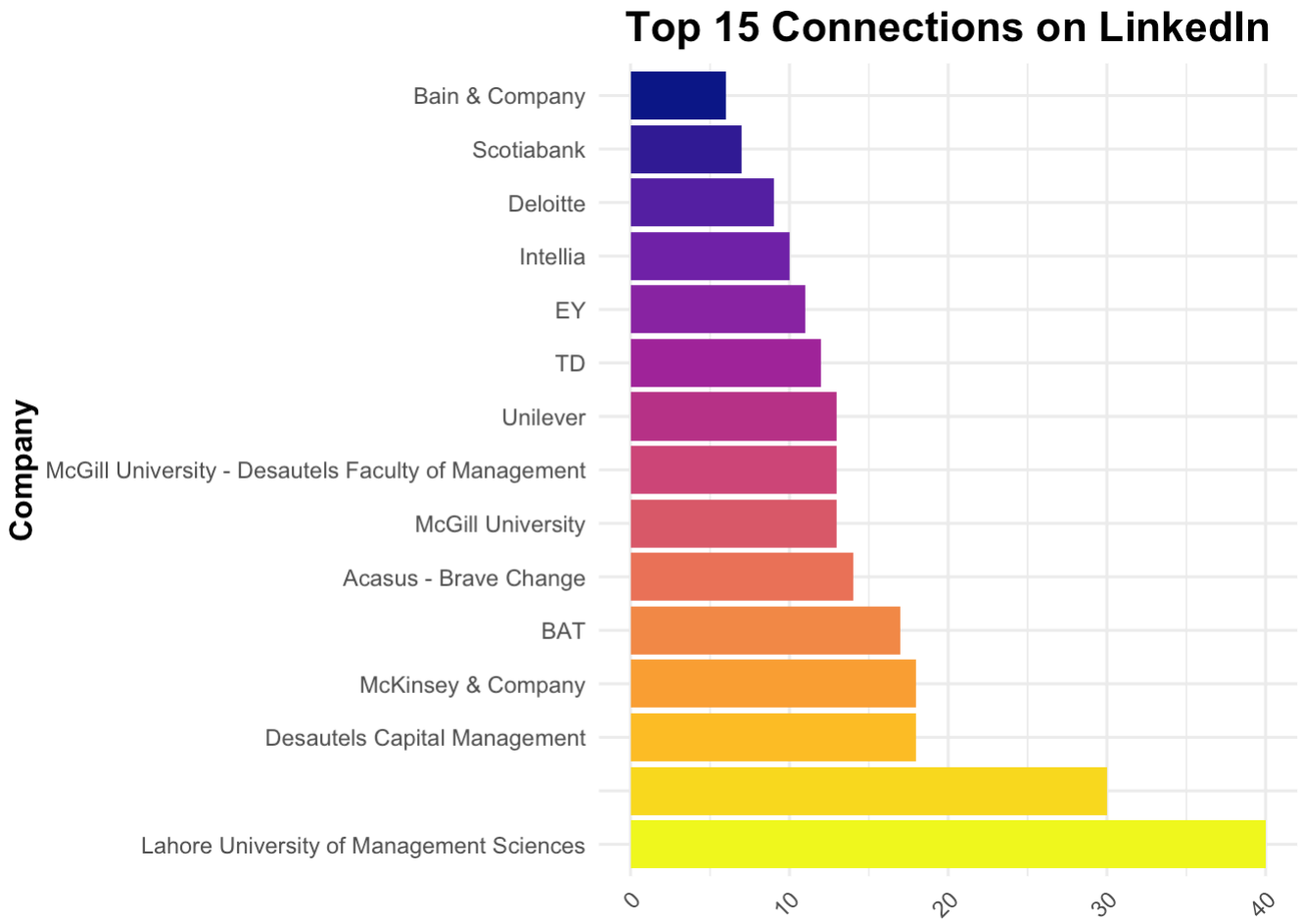
Frequency Table and Bar Chart

```
connections <- connections %>%
  select(Name = `First.Name`, Company, Position)
freq_table <- table(connections$Company)
freq_table <- sort(freq_table, decreasing = TRUE)
top15 <- head(freq_table, n = 15)
knitr::kable(as.data.frame(top15), col.names = c("Company", "Connections"))
```

Company	Connections
Lahore University of Management Sciences	40
	30

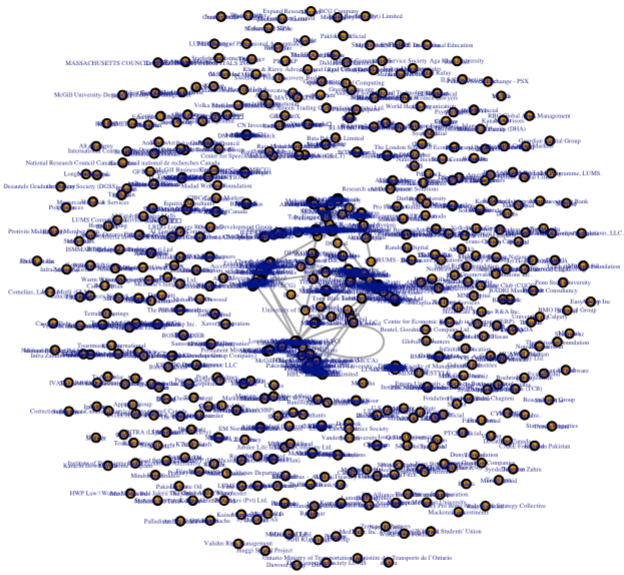
Company	Connections
Desautels Capital Management	18
McKinsey & Company	18
BAT	17
Acasus - Brave Change	14
McGill University	13
McGill University - Desautels Faculty of Management	13
Unilever	13
TD	12
EY	11
Intellia	10
Deloitte	9
Scotiabank	7
Bain & Company	6

```
top15_df <- as.data.frame(top15)
names(top15_df) <- c("Company", "Connections")
top15_df$Company <- factor(top15_df$Company, levels = top15_df$Company)

ggplot(top15_df, aes(x = reorder(Company, -Connections), y = Connections, fill = Comp
any)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_viridis_d(direction = -1, option = "C") +
  labs(title = "Top 15 Connections on LinkedIn",
       x = "Company",
       y = "Number of Connections",
       fill = "Company") +
  theme_minimal() +
  theme(legend.position = "none",
        plot.title = element_text(size = 16, face = "bold", hjust = 0.5),
        axis.text.x = element_text(angle = 45, hjust = 1),
        axis.title.x = element_blank(),
        axis.title.y = element_text(size = 12, face = "bold"))
```



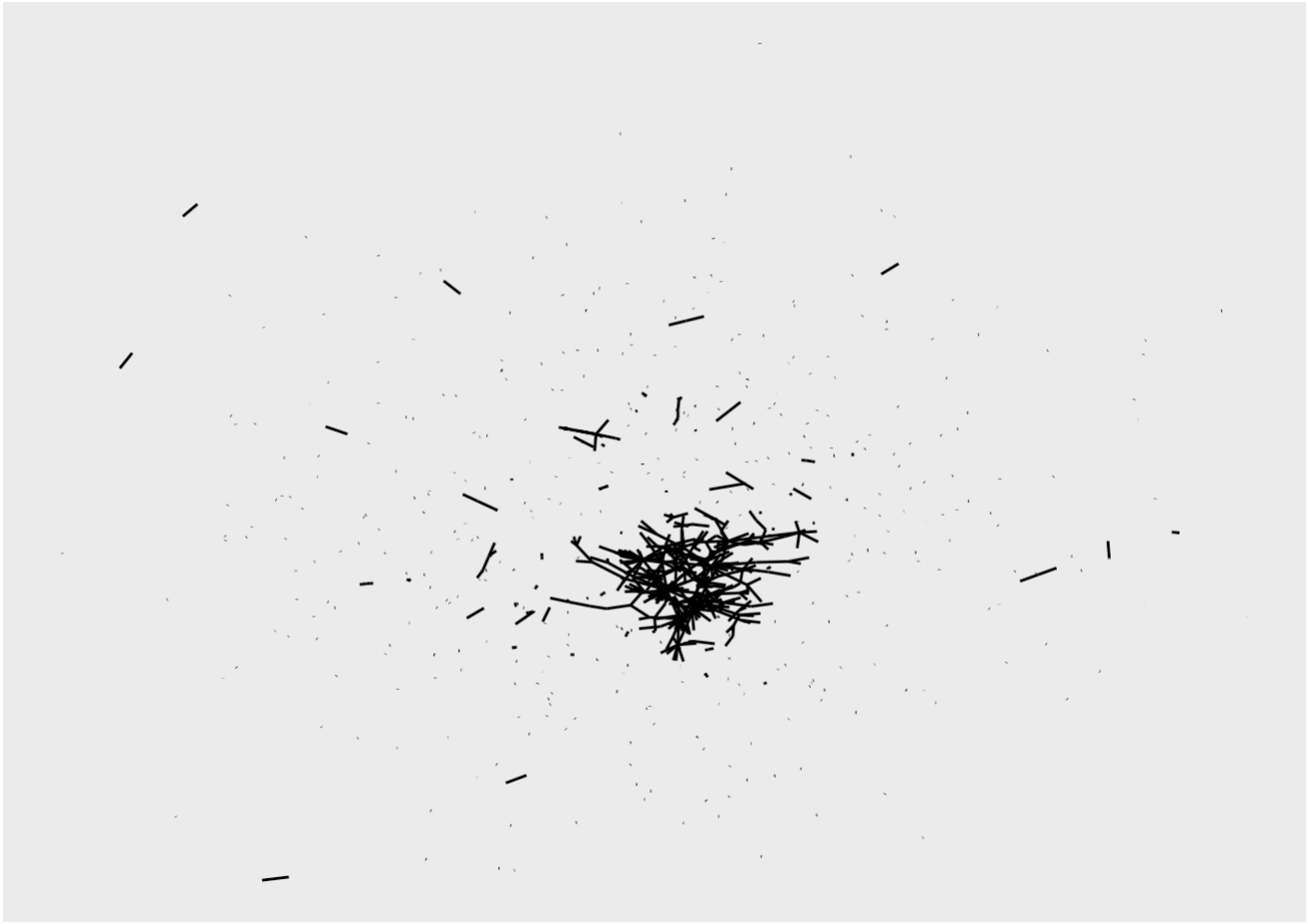
Network Creation and Visualization

```
people <- connections %>% distinct(Name) %>% rename(label = Name)
companies <- connections %>% distinct(Company) %>% rename(label = Company)
nodes <- full_join(people, companies, by = "label") %>% rowid_to_column("id")
edges <- connections %>%
  select(Name, Company) %>%
  left_join(nodes, by = c("Name" = "label")) %>%
  rename(from = id) %>%
  left_join(nodes, by = c("Company" = "label")) %>%
  rename(to = id) %>%
  distinct(from, to)
network_igraph <- graph_from_data_frame(d = edges, vertices = nodes, directed = TRUE)

plot(network_igraph, vertex.size = 3, vertex.label.cex = 0.2, edge.arrow.size = 0.01)
```



```
tbl_network <- as_tbl_graph(network_igraph)
ggraph(tbl_network, layout = "fr") + geom_edge_link()
```

```
## Warning: Using the `size` aesthetic in this geom was deprecated in ggplot2 3.4.0.
## ℹ Please use `linewidth` in the `default_aes` field and elsewhere instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



Interactive Network Visualization with visNetwork

```
nodes$color <- ifelse(nodes$label == "McGill University - Desautels Faculty of Manage
ment", "red", "gray")
nodes_vis <- nodes %>% select(id, label, color) %>% mutate(title = label)
edges_vis <- edges %>% select(from, to)
visNetwork(nodes_vis, edges_vis) %>%
  visNodes(color = list(background = nodes_vis$color, border = "#2b2b2b")) %>%
  visEdges(arrows = 'to') %>%
  visOptions(highlightNearest = TRUE, nodesIdSelection = TRUE)
```

Select by id